DOCUMENT RESUME

ED 452 219 TM 032 504

AUTHOR	Walker, Sherry Freeland, Ed.
TITLE	High-Stakes Testing: Too Much? Too Soon?
INSTITUTION	Education Commission of the States, Denver, CO.
ISSN	ISSN-0736-7511
PUB DATE	2000-00-00
NOTE	22p.; Theme Issue. Published three times a year. ECS Distribution Center, Education Commission of the States, 707 17th Street, Suite 2700, Denver, CO 80202-3427 (Annual subscription, $20). Tel: 303-299-3600; For full text: http;//www.ecs.org.
PUB TYPE	Collected Works - Serials (022)
JOURNAL CIT	State Education Leader; v18 n1 Win 2000
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	*Accountability; *Educational Testing; Elementary Secondary Education; *High Stakes Tests; Scores; *State Programs; Test Results; *Test Use; Testing Programs

ABSTRACT

This theme issue focuses on the use and consequences of high stakes tests. The lead article, "High-Stakes Testing: Too Much? Too Soon?" by Sherry Freeland Walker, introduces the topic and related issues, outlining the pros and cons of high stakes testing by the states. The problem, some experts say, is that states have tried to do too much too soon without the proper preparation and support for everyone involved. "The History of Testing," by Sherry Freeland Walker, traces the growth of high stakes testing through the last half century and in the present context of the standards movement. "High-Stakes Assessments Bring Out the Critics," by Jennifer Dounay, discusses a number of criticisms of high stakes testing and some responses from the public. "Why Is 'Teaching the Test' a Bad Thing?" by Lorrie Shepard, explores the issues of test score inflation, curriculum distortion, and safeguards against political pressures in testing. "How States Are Responding to Low-Performing Schools," by Katy Anthes, Susie Saavedra, Judie Mathers, and Jane Armstrong, describes the interventions states with high stakes accountability systems are using with low performing schools. The effects of high stakes tests on teacher education are outlined in "High-Stakes Testing Pressures Teacher Education" by Michael Allen. Other articles in this issue are: (1) "Maryland Moves toward Intervention" (Mary Fulton); (2) "Texas Test Withstanding Court Scrutiny" (Jill Weitz); (3) "Why Do We Need High-Stakes Assessments?" (Michael Sentance); (4) "Poor Test Results Lead to Math Consortium"; and (5) "Performance Management, Not Just Accountability" (Peter Robertson). (SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 452 219

State Education *leader*

Volume 18 Number 1 Winter 2000

High-Stakes Testing: Too Much? Too Soon?

TM032504

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

S. F. Walker

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

Education Commission of the States

BEST COPY AVAILABLE

TM

STATE EDUCATION Leader

EDUCATION COMMISSION OF THE STATES

ECS

VOLUME 18 NUMBER 1 WINTER 2000

HIGH-STAKES TESTING: TOO MUCH? TOO SOON?

The mission of the Education Commission of the States (ECS) is to help state leaders identify, develop and implement public policy for education that addresses current and future needs of a learning society.

by Sherry Freeland Walker

Pick up the newspaper and you're likely to see an article about state assessments that carry big consequences. Parents and students in one state protest the test, while policymakers in another laud or bemoan the outcomes of their state's latest assessment. Across the nation, state and district leaders are putting more emphasis on testing and using test results to make more decisions about students and schools. Will this student be promoted to the next grade? Will that one graduate from high school? Should this school be reconstituted?

Using assessment tests for such "high-stakes" purposes is gaining public support as a way to determine how good a job public schools are doing. Policymakers see them as a way to raise standards and achievement and hold students and educators accountable. But as support grows on one hand, so does opposition on the other. Are high-stakes tests worthwhile? Or is the controversy around them likely to derail the standards movement?

Lagging skills

No one disputes that too many American students are not gaining the knowledge and skills they need to succeed in college and the workforce. Only about one-third are proficient in reading and fewer still in math, according to National Assessment of Educational Progress

Continued on next page

In this issue

■ [illegible]
■ [illegible]
■ [illegible]
■ [illegible]

STATE EDUCATION LEADER

Published three times a year. $20 per year.

Letters, manuscript queries and unsolicited submissions should be addressed to the editor. No payment for articles.

© Education Commission of the States, 2000. All rights reserved.

ECS is a national, nonprofit organization that helps governors, state legislators, state education officials and others develop policies to improve the quality of education at all levels. ECS offices are located at 707 17th Street, Suite 2700, Denver, Colorado 80202-3427. 303.299.3600. FAX 303.296.8332. e-mail: ecs@ecs.org. http://www.ecs.org.

To request permission to excerpt part of this publication, call or print or electronically, please fax to ECS Publications at 303.296.8332.

Articles do not necessarily reflect views of ECS.

It is ECS policy to take affirmative action to prevent discrimination in its policies, programs and employment practices.

Continued from previous page

scores. Even the most advanced U.S. students lag behind their peers in other countries on the Third International Mathematics and Science Study. And public opinion shows Americans increasingly critical of public schools overall.

Almost every state has set content standards for what students should know and measuring whether students are meeting those standards is a natural outgrowth. To date, 17 states, the District of Columbia and Puerto Rico have policies that base promotion or retention on a student's score on a state and/or district assessment (see page 11). Twenty-seven states have high school exit exams (though not all are tied to graduation or test beyond 9th-grade skills). Polls consistently show public support for standardized testing.

Pros and cons

Proponents of high-stakes testing argue that it leads to achievement and other gains:

- Students know what is expected and that the test really counts, so they work harder.
- Schools identify and can address student weaknesses early.
- Similarly, schools discover areas of overall weakness, prompting them to refocus resources where they are most needed.
- Education across the state is more consistent, eliminating situations where schools in some districts are superior to others.
- The public sees gains from year to year and regains confidence in public schools.

Critics say the tests sometimes are too hard, lead teachers to teach to the test, take time away from instruction, and are expensive. Teachers say they're unprepared to teach to the standards, and students claim they're being tested unfairly, on content they haven't yet had. Some parents and students are calling for an end to high-stakes testing, and some policymakers are reexamining plans to tie tests to key decisions such as graduation or to make high-stakes tests the central part of an accountability system (see pages 4-6).

Too much, too soon?

The problem, some experts say, is that states have tried to do much too soon without the proper preparation and support for everyone involved.

"Teachers and principals simply do not know how to do what they are expected to do with the new standards," said Richard F. Elmore, Harvard School of Education professor, at a recent Washington, D.C., conference.

While some policymakers are rethinking assessments, others say the low scores are just an indication of the work that needs to be done. "When we fired this missile," Todd Bankofier of the Arizona Board of Education said, "we knew we had to guide it. It's going to take some left turns and some right turns, but it would be wrong to turn it completely back."

"Doing away with the tests or the consequences is the easy way out," Robert Schwartz and Matthew Gandal wrote in the January 19, 2000, issue of *Education Week*. "It allows us to avoid the hard work of improving instruction and restructuring the use of time and resources so that all students are given the time and support needed to meet standards."

Confronting the dilemma

Jay P. Heubert and Robert M. Hauser of the National Research Council's Committee on Appropriate Test Use recommend in *High-Stakes Testing for Tracking, Promotion and Graduation* that policymakers keep the following principles of appropriate test use in mind:

- **Use the right test.** Tests are valid only when used for the specific purpose for which they were designed.
- **Remember tests are not perfect.** Questions are but a sample of possible questions that could be asked in a given area.
- **Don't use a test as the sole determinant of a major decision.** Promotion and graduation decisions should be based on many factors.
- **Don't justify bad decisions with a test score or any other kind of information.** Tests will not lead to better outcomes if districts and schools lack the services to help students who don't come up to standard.

The answer to who's right — the critics or the supporters — seems to be both. If the right test is used in the right way, in conjunction with other measurements, it can be an effective way to assess student learning. Without attention to factors such as discrimination, curriculum and accuracy, however, it can be detrimental to both students and schools alike.

This issue of *State Education Leader* looks at the controversy around high-stakes testing.

Freeland Walker is ECS publications director.

BEST COPY AVAILABLE

THE HISTORY OF TESTING

by Sherry Freeland Walker

> "... [F]ormal testing has become the kudzu of modern American society, a healthy vigorous grower penetrating all available space."
>
> — Gary Natriello, Columbia University's Teachers College, and Aaron M. Pallas, Michigan State University

Testing is big news these days, and the stakes are getting higher and higher. As business and the public put more pressure on public schools and students to achieve at higher levels, the use of testing is expanding rapidly.

Throughout the last century, the uses of standardized testing, and the reasons for using it, have grown considerably. As Gary Natriello of Columbia University's Teachers College and Aaron M. Pallas of Michigan State University, say, "formal testing has become the kudzu of modern American society, a healthy vigorous grower penetrating all available space."

Half a century

Standardized testing has been a feature of public schools for half a century, initially serving largely to compare schools and students against a standard set by testing companies. Another use was to "sort" students, such as identifying those considered fit for higher education versus those who would be better suited to vocational school.

The 1970s saw an eruption of interest in "minimum competency testing." Then, as now, say Robert Linn and Joan Herman of the National Center for Research on Evaluation, Standards and Student Testing, "reformers sought to improve education by holding educators and students accountable for achieving standards of performance, using tests for high school graduation and or grade-to-grade promotion."

By the early 1980s, nearly three-fourths of the states had minimum competency testing requirements. Most took the form of multiple-choice items that students either passed or failed and primarily pinpointed gains at the low end of the spectrum. The tests did little if anything to measure how much students were learning or how advanced their skills were.

Standards movement

Growing criticism of public schools led policymakers and educators to turn toward testing to measure higher skills and to gain support for raising standards. The late 1980s saw the rise of assessment tied to accountability for student and school performance, although states were relying heavily on nationally published standardized tests, rather than assessments geared to individual state standards.

The early days of test results tied to accountability, however, were criticized as showing an inflated pattern of scores. Because the tests suddenly had high stakes, teachers were teaching to the test, critics said. They based their reasoning largely on the fact that gains on the National Assessment of Educational Progress tests were not as high as scores on other assessments.

While the current wave of education reform continues to emphasize accountability, it is more tied to the setting and implementing of state standards, both content (what students should know) and performance (how well they are able to do it). States are aligning assessments to their standards and demanding much more from students than they have previously.

Freeland Walker is ECS publications director.

Education Commission of the States
STATE EDUCATION LEADER
VOL. 18 NO. 1 WINTER 2000

HIGH-STAKES ASSESSMENTS BRING OUT THE CRITICS

by Jennifer Dounay

> "As the assessment stakes have increased for both students and schools, various 'stress points' in the system are causing some students, parents and others to question the validity of assessment and accountability systems."

For many people, both in and outside the education policymaking field, the concept of assessing students on their knowledge and skills seems a perfectly innocuous proposition. After all, why shouldn't pupils be held accountable for learning what they have been taught during a given school year or by a certain milestone in their school careers?

This proposition, however, is not as simple as it may appear. As the assessment stakes have increased for both students and schools, various "stress points" in the system are causing some students, parents and others to question the validity of assessment and accountability systems.

Too much pressure

Parents in some states are asserting that some high-stakes tests place undue pressure on young children. Stories of increasing numbers of children suffering from sleep disorders and other stress-related maladies have appeared in the press in the past few years.

Districts across the nation have offered Saturday and summer tutorial classes to give children extra time to work on skills that may be tested. The Hartford, Connecticut, schools offered classes during the 1999 spring break to help 3rd, 5th and 7th graders prepare for the Connecticut Mastery Test scheduled for the fall. (To the district's credit, scores did improve significantly.)

Kaplan, known for its SAT and ACT preparation books, has released books to help students and parents of young children prepare for standardized tests in Florida, New York, Texas and Massachusetts.

"Dumbing down" of the curriculum

Another criticism is that the curriculum may be "dumbed down" as a result of state-mandated testing. Some people fear rote memorization may be stressed rather than problem-solving skills and that teachers will focus on subject areas or facts most likely to appear on assessments, rather than more complex skills, such as critical thinking.

There also is widespread concern that subjects not tested (for instance, fine arts or physical education) will be accorded less class time or set aside altogether, as some elementary schools have done with recess, to spend more time on academics.

Critics also argue that too much time is taken away from instruction when students are coached on testing techniques and then spend hours taking the tests.

Score discrepancies

Parents, as well as the general public, also doubt the integrity of a state assessment when scores do not match their children's grades or achievement measured by other tests. Numerous media articles have profiled students with

"A" or "B" averages who attain low scores on state assessments or fail to pass high school exit examinations. Parents and students wonder whether grades are inflated or if the bar on the state assessments has been set unreasonably high. Parents in affluent areas of New York such as Rye, Great Neck and Mamaroneck were shocked, for example, when, according to a November 1999 *New York Times* article, one in five of their children failed the state's new 8th-grade math assessments.

State issues

Discrepancies between indicators of student achievement have shown up at the state and district levels as well. For example, Virginia began in 1998 to assess 3rd, 5th and 8th graders as well as high schoolers on the state's Standards of Learning (SOLs) in English, history/social sciences, mathematics and science. Starting in the 2006-07 academic year, only schools whose pass rates meet or exceed 70% in the four subject areas will be eligible for accreditation, with the exception of 3rd-grade science and history, whose minimum pass rate for accreditation will be 50%.

Results of the spring 1999 tests reveal much work to be done — only 6.5% of Virginia schools met the pass-rate standard in all four of the subjects. In Fairfax County, where students posted an average SAT score of 1095 in 1998 (versus a national average of 1005) and where 91% of students continue to postsecondary education, only 54% passed the SOLs in 1998.

Because of these discrepancies, Virginia has taken measures to evaluate the fairness of the SOLs assessments. In February 1999, testing experts from three universities declared the SOLs valid and reliable. And a new SOLs Test Technical Advisory Committee has been commissioned to report annually on the assessments' validity and reliability and propose suggestions and recommendations for future changes.

Massachusetts' assessment results likewise have raised eyebrows in that state. The Massachusetts Comprehensive Assessment System (MCAS) tests 4th, 8th and 10th graders in English language arts, math and science and technology. In September 1999, the State Board of Education voted to rate schools in two-year cycles based on their students' performance. Schools that do poorly must submit improvement goals to the state which, if unmet in two years, will open the schools to state takeover.

Individual students likewise will feel the effect of the MCAS. The class of 2003 will be the first whose high school graduation will depend upon students' scoring at the proficient or advanced level on all of the 10th-grade tests. Like Virginia, scores so far have been low. In 1999, only 34% of students reached those levels in English language arts, 24% in mathematics, and 24% in science and technology.

Minority discrimination

Some test critics point out that students from predominantly white and middle- to upper-class districts score the highest on high-stakes and other assessments. An analysis of the 1998 MCAS tests, conducted by the Gaston Institute for Latino Community Development at the University of Massachusetts-Boston, found that cities with the highest proportions of Hispanic test takers fared worst on the 10th-grade math assessments, with failure rates nearly as high for African-American students. While the statewide average failure rate for students of all races on this assessment was 52%, it was 83% for Hispanic students and 80% for African-American students.

Testing programs in other states have turned up similar gaps in minority achievement, although Texas' system — the Texas Assessment of Academic Skills (TAAS) — recently survived a legal challenge that claimed the high school exit exam discriminates against Hispanics and blacks (see page 10 for more). While recognizing the differences in passage rates among blacks (60%), Hispanics (64%) and whites (86%) in the spring 1999 administration, U.S. District Judge Ed Prado wrote:

> "The evidence suggests that the State of Texas was aware of probable disparities and that it designed the TAAS accountability system to reflect an insistence on standards and educational policies that are uniform from school to school."

Mistakes and cheating

High-visibility examples of security breaches, teacher and administrator cheating, and mistakes made by testing companies also have shaken the public's confidence in assessment systems.

Essay questions for Ohio's 4th- and 8th-grade writing assessments had to be rewritten after a paper quoted students discussing the essay questions before some schools in the state had administered them. Rhode Island

Continued on next page

"Students from predominantly white and middle- to upper-class districts score the highest on high-stakes and other assessments."

Continued from previous page

postponed administering mathematics and English assessments for 4th, 8th and 10th graders last year after widespread security breaches were discovered.

Test-tampering cases in Houston, Austin and eight other Texas districts may have been the impetus for the September 1999 creation of the state's Public Education Integrity Task Force. In New York City, 52 teachers and administrators were named in a December 1999 report for helping students improve their test scores by a variety of means.

Mistakes in scoring also have occurred. Writing assessments for 4th, 7th and 10th graders in Washington State were rescored by hand and subsequently released two months behind schedule after scoring mistakes were discovered in summer 1999. In September 1999, testing company CTB/McGraw Hill informed officials in Indiana, North Carolina, South Carolina, Wisconsin and New York City that their tests may have been scored incorrectly. Ramifications of the blunder were especially strong in New York City, where more than 8,600 students were erroneously placed in summer school as a result of "low" test scores.

Backlash

Such cases of confusion, potential unfairness and frustration have led to public outcry against tests in some locales and responses from decisionmakers. The results of the math portion of Arizona's new assessment instrument, which members of the class of 2002 must pass to graduate, revealed that 0% of the 44,245 students who took the test exceeded the standard in math and only 11% met the standard.

In response to cries from parents, students and educators across the state that the test is too difficult, the state board agreed to reexamine the scoring levels. Likewise, the Virginia state board has indicated it is open to discussion of changing the history portion of the SOLs, on which significantly fewer students attain the proficient level than in other subjects that the state tests.

Isolated instances of civil disobedience as well as organized resistance to high-stakes assessments have appeared in several states. Students in some Massachusetts cities sat out the spring 1999 administration of the MCAS. A teacher in Harwich refused to give his students the 8th-grade history test after noticing that some questions dealt with the Civil War, which students had not yet studied. Groups such as the Coalition for Authentic Reform in Education and Cambridge Parents Against the MCAS have been established. Parents in several cities, including Boston, "have encouraged their children to boycott the test or have taken them out of the public schools," according to an October 31, 1999, *Boston Globe* article.

Likewise, the *Christian Science Monitor* reported that "in certain Detroit suburbs — particularly Birmingham, Troy and Farmington — protesting parents have refused to allow their children to take the state test. In some towns, fewer than 15% of students participated in state testing — a number so small as to render any results meaningless." The same article notes that students intentionally have failed tests or refused to take them in California, Wisconsin and Illinois as well.

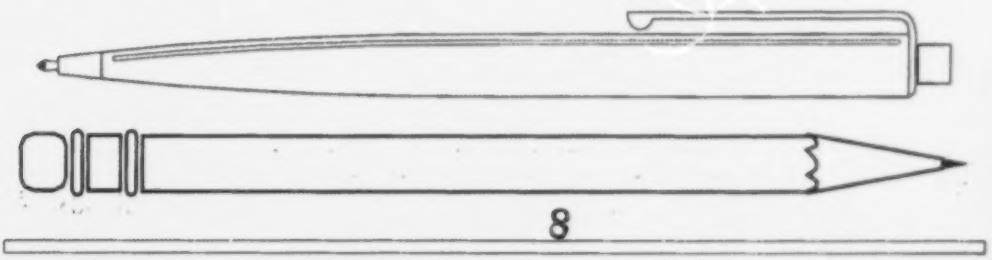
What's next?

What's a policymaker to do? After all, testing experts themselves caution that when higher standards and new assessments are implemented, scores will reflect the greater challenges placed upon students and the teachers who must prepare them.

There are no simple solutions. Policymakers, however, must be cautious to avoid alienating their constituencies or dismissing parents' concerns. Above all, they must remember that, while scores may reflect improvements in schools or the tests themselves, the final goal of states' standards and assessment systems is not necessarily the race for ever-higher scores but the race for students' solid preparation for the workplace or postsecondary education.

Dounay is an ECS research associate.□

"[Policymakers] must remember that, while scores may reflect improvements in schools or the tests themselves, the final goal of states' standards and assessment systems is not necessarily the race for ever-higher scores but the race for students' solid preparation for the workplace or postsecondary education."



by Lorrie Shepard

WHY IS "TEACHING THE TEST" A BAD THING?

According to a recent survey reported by *Education Week*, testing is the number one accountability tool, adopted in 48 of 50 states. Test results are intended to focus attention on raising student achievement. Yet, critics complain that the emphasis on testing leads to problems of "teaching the test." What is meant by that, and why is it a bad thing?

Typically, teaching the test means devoting extended time to subject areas that are tested, such as reading and math, to the exclusion of other subjects. Test format becomes a template for how tested subjects are taught. Worksheets and practice assessments mirror the anticipated accountability tests as much as possible. A recent study in Texas, for example, found that teachers in urban schools were required to use test-prep materials from September through March, when the Texas Assessment of Academic Skills test was given.

Test-score inflation

When tests are developed initially, they are designed to reflect curriculum frameworks or content standards. Particular test questions are intended only to be samples of the full curriculum. How students do on the test is supposed to show how well they have mastered that curriculum. But if students practice only questions that imitate the test, test performance may no longer "generalize" to the intended curriculum content. In fact, controlled studies have shown that students may not be able to answer the same questions if asked even in slightly different ways.

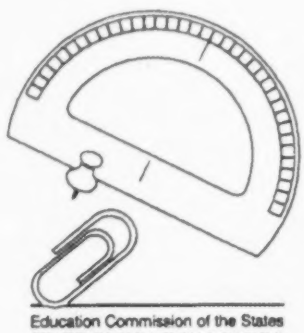
In one classic experimental study, all students in a study were taught to translate from Roman to Arabic numerals. The group tested in the same order did well, but when the other group was asked to translate in reverse — from Arabic to Roman numerals — the drop-off in performance was startling. Students lost from 35 to 50 percentile points, showing they never understood how the number system really works.

Curriculum distortion

The negative effects of teaching the test on student learning are the flip side of test-score inflation. In a nationwide survey for the National Science Foundation, the majority of teachers acknowledged shifting instructional emphasis from nontested to tested topics and, at the same time, reported negative impacts of mandated testing on curriculum and learning. Although critics originally feared that testing would take instructional time away from "frills," such as art and citizenship, research shows that untested subjects such as social studies and science have been relegated to Friday afternoons or even eliminated.

Continued on next page

"[A] test-driven curriculum encourages teaching of skills in isolation, which may deny students the very activities that might have made the problems understandable and useful."



> **"The movement toward performance assessments is aimed at correcting the distorting effects of multiple-choice test formats."**

Continued from previous page

Even in tested subjects, instruction is focused only on skills covered by the test. In a study by Mary Lee Smith, elementary teachers had given up reading real books, writing and long-term projects and were focusing on word recognition, recognition of spelling errors, language usage, punctuation and arithmetic operations.

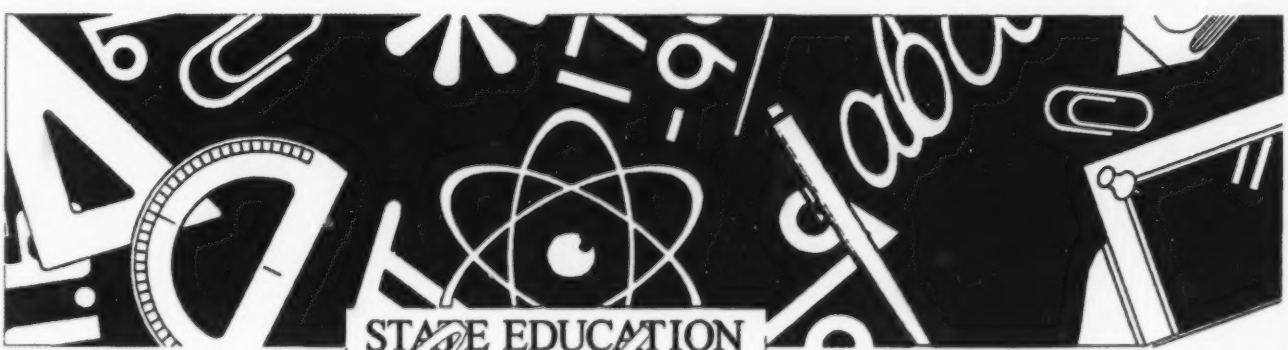
Unfortunately, a test-driven curriculum encourages teaching of skills in isolation, which may deny students the very activities that might have made the problems understandable and useful. Practicing only test-like formats also elicits different cognitive processes than working with more extended and challenging curricular materials. For example, students are asked to read artificially short passages and search for answers to formulaic questions. They practice finding mistakes rather than doing significant writing on their own, and they learn to guess by eliminating wrong answers.

Safeguards

Developing new forms of the test each year is one limited safeguard that prevents practicing on specific test items. In addition, the movement toward performance assessments is aimed at correcting the distorting effects of multiple-choice test formats. The more that extended tasks on tests reflect the actual kinds of written expression, problem solving and applications of knowledge that are intended in the curriculum, the less likely it is that teaching to the test will distort either learning or test-score gains.

The content of a test alone, however, cannot be sufficient safeguard against political pressures. Ultimately, the best remedies are (1) to put less weight on a single indicator when judging the quality of schools and (2) acknowledge accurately that the responsibility for student achievement is shared among students, parents, teachers, school administrators, community leaders and policymakers.

Shepard is professor of education, University of Colorado at Boulder.



STATE EDUCATION Leader

Published three times a year. Annual subscriptions are $20. Subscribe now.

Name _____

Address _____

_____

City, State, ZIP _____

Telephone ( ) _____

FAX and e-mail _____

Payment Method: ❑ Check enclosed — Payable to Education Commission of the States

Credit card: ❑ MasterCard ❑ Visa ❑ American Express

Credit card account # _____ Expiration date _____

Cardholder's signature _____

Mail check/order form to: ECS Distribution Center, Education Commission of the States, 707 17th Street, Suite 2700, Denver, CO 80202-3427

FAX credit card orders to: 303-296-8332, e-mail jivey@ecs.org, or call 303-299-3692.

MARYLAND MOVES TOWARD INTERVENTION

by Mary Fulton

Maryland has been a pioneer state in developing academic standards and assessments that measure student progress toward the standards. But lower-than-expected scores and a new high school graduation assessment are leading state officials to take steps to ensure all students meet the rigorous accountability requirements.

In the early 1990s, the state developed a criterion-referenced test to measure school performance and progress toward meeting state standards. Students in grades 3, 5 and 8 take the Maryland School Performance Assessment Program (MSPAP) in reading, writing, language usage, mathematics, science and social studies. The test is designed to provide information to improve instruction and measure school improvement, not individual student performance. Each elementary and middle school was to have 70% of its students scoring at the satisfactory level by this year.

Scores for 1999 showed these results:

- ☐ Statewide, 43.8% of students are scoring at the satisfactory level (up 12 percentage points since 1993), when scores are averaged across the various tests.
- ☐ In 77 of 1,357 schools, at least 70% of students scored at the satisfactory level, up from 11 schools in 1993.
- ☐ Twenty of the state's 24 school districts averaged 40% or more students at the satisfactory level, up from four districts in 1993.

Overall, MSPAP results have been mixed. While several schools and districts have made gains — some significant — many schools have seen modest increases or fluctuating scores. In January 2000, state officials announced that they are considering taking over 10 Baltimore schools with consistently low performance on the assessments and other measures.

High school assessment

The new high school assessment (which, unlike the current 9th-grade assessment, is linked to standards) will be field-tested this spring. From this pilot phase, the State Board of Education will determine the number of tests required for graduation and define the passing rates. The class of 2005 will be the first to take the exams, which will gauge individual student, as well as school, performance.

This assessment will include 12 end-of-course tests in English, mathematics, science and social studies. Students must pass three tests – English, algebra or geometry, and government – to graduate. Local districts have discretion to require a biology test as well. The state board will determine when additional tests should be implemented.

Intervention and prevention

Failure of schools to meet the 70% passing goal on the MSPAP, and introduction of the new high school test, have led state board members to realize that many Maryland students lack the necessary preparation to pass the assessments. In October 1999, the state board approved an initiative, "Every Child Achieving: A Plan for Meeting the Needs of Individual Learners," focusing on academic intervention, educator and administrator capacity, and student readiness. If fully funded, the initiative will require:

- ☐ Extended-learning experiences (before and after school, on Saturdays, etc.) for K-8 students with deficiencies in reading and math.
- ☐ Summer program for students not reaching proficiency levels in reading and/or mathematics by the end of grade 8. Students who don't reach proficiency levels will be allowed to enroll in high school, but not in core courses until they reach required levels.
- ☐ Individualized learning plans for students who fail one or more high school assessments.
- ☐ Newly hired elementary teachers to have strong content knowledge in core subject areas.
- ☐ Newly hired secondary teachers to have a major in content area they will teach.

The initiative is one of the first comprehensive intervention/prevention plans explicitly tied to a state's high-stakes assessment. The focus on intervention, teacher and administrator professional development, and early childhood education should provide a more solid foundation as the stakes are increased for Maryland students.

Fulton is an ECS policy analyst.☐

"[Maryland's] initiative is one of the first comprehensive intervention/ prevention plans explicitly tied to a state's high-stakes assessment."

Continued on next page

BEST COPY AVAILABLE

TEXAS TEST WITHSTANDS COURT SCRUTINY

by Jill Weitz

Texas' high-stakes assessment, which students must pass to receive a diploma, recently survived a court challenge that claimed the test harms minority students. Plaintiffs argued that using the Texas Assessment of Academic Skills (TAAS) to determine who can graduate violates federal civil rights and due process laws. The federal court, however, found that the disparity among white, African-American and Hispanic students' test scores is a reasonable step on the road to increased achievement for all.

The decision stated:

> "[The] court has had to weigh what appears to be a significant discrepancy in pass scores on the TAAS test with the overwhelming evidence that the discrepancy is rapidly improving and that the lot of Texas' minority students, at least as demonstrated by academic achievement, while far from perfect, is better than that of minority students in other parts of the country and appears to be getting better."

While the court acknowledged the harm to minority students who drop out or are refused a diploma, it rejected the idea that these circumstances were sufficient to overcome the state's interest in improving the education system as a whole.

> "State officials say the testing program is making schools focus more on academic achievement, although opponents argue that students who fail the high school test are simply dropping out."

Accountability system

Since the trial began last September, states have been poised to see if the court's decision would change the legal precedent that generally has upheld high-stakes exit exams against claims of racial discrimination. Texas' high-stakes testing system, which measures student performance toward academic goals and is the basis for the state's accountability system, requires high school students to pass TAAS or end-of-course exams in specified subjects. Students have eight opportunities to take the test before graduation and may take remedial courses in any areas they fail.

TAAS assessments are given in early grades as well and are, or soon will be, the basis for promotion or retention. Students who fail will receive accelerated instruction and at least two additional opportunities to take the test.

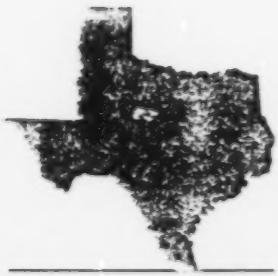
Achievement up

Test data from the Texas Education Agency show that TAAS achievement levels increased from spring 1994 to spring 1999. The percentage of students in grades 3-8 and 10 passing the test (scoring 70% or higher) rose from 53% to 78%. Students meeting minimum requirements on the reading, math and writing tests rose by 12%, 28% and 12%, respectively. State officials say the testing program is making schools focus more on academic achievement, although opponents argue that students who fail the high school test are simply dropping out.

Proponents' arguments are bolstered by Texas' increased achievement levels on the National Assessment of Educational Progress, as well. And last month, the National Educational Goals Panel recognized Texas as one of only 12 states that has made great progress toward achieving the national education goals and cited Texas for its improvement in student performance.

"This ruling keeps our testing program and accountability system intact, which I believe is good for Texas," Commissioner of Education Jim Nelson said.

Weitz is a former ECS policy analyst.



- For more information on the legal implications of high-stakes assessments, see *High-Stakes Testing for Tracking, Promotion and Graduation*, available from the National Academy press at books.nap.edu/catalog/6336.html.
- See www.txwd.uscourts.gov for the full text of the Texas court decision.
- See www.tea.state.tx.us/student.assessment/ for more information on the Texas assessment system.
- For test results, see the Texas Education Agency Web site at www.tea.state.tx.us/student.assessment/results/swresult/g310allau99.htm.

WHY DO WE NEED HIGH-STAKES ASSESSMENTS?

by Michael Sentance

The simple answer to this question is that we need standards for teachers and students because, under the old system, too many students fail to learn the challenging curriculum they need and deserve. In addition, the old assessment tools — nationally standardized or college entrance tests — do not measure how well students can meet the demands of the workforce or the rigors of university-level work.

The first question, of course, that a parent asks in a school conference is: "How is my child doing?" For years, the teacher could assuage parents' concerns through a review of the child's grades, the results of nationally normed standardized tests and classroom observations. And when the student didn't get into his or her first choice of college or was denied a job, he was usually blamed for the outcomes — he didn't work hard enough, or she had no aptitude for math.

In Massachusetts, we tried standards and assessments — but without consequences. For a decade, the results remained largely unchanged as local educators professed surprise with poor scores. It was not until consequences were attached to the tests that we began to give parents a more accurate picture of how their district, their school — and their child was doing.

Sentance is the education policy advisor to the governor of Massachusetts and an ECS commissioner.

States That Base Promotion and Retention on State and/or District Assessment

Arizona	Florida	North Carolina
Arkansas	Illinois	Oklahoma
California	Louisiana	South Carolina
Colorado	Michigan	Texas
Connecticut	Mississippi	Wisconsin
Delaware	New Mexico	

Source: "State Student Promotion/Retention Policies," ECS Web site, www.ecs.org.

States with High School Graduation Exit Examinations*

In mid-1999, 26 states and Puerto Rico had high school exit examinations. Use of the data collected from the tests ranges from determining which students need remediation to which students will graduate. Some states use the data to develop improvement plans, publish state report cards, assess school weaknesses, direct curriculum improvements, and/or evaluate staffing and resources.

Alabama	Maryland	Ohio
Alaska	Massachusetts	Puerto Rico
California	Minnesota	South Carolina
Delaware	Mississippi	Tennessee
Florida	Nevada	Texas
Georgia	New Jersey	Utah
Hawaii	New Mexico	Virginia
Indiana	New York	Washington
Louisiana	North Carolina	Wisconsin

*as of July 31, 1999

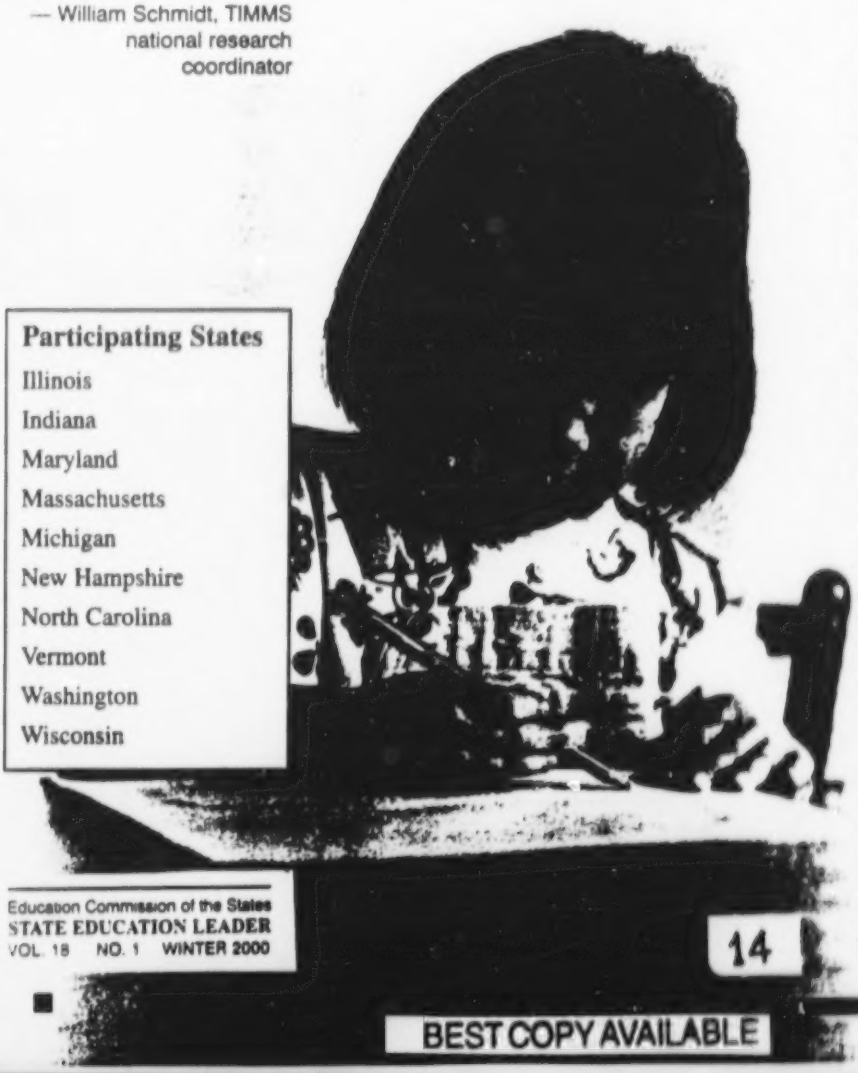
Source: National Governors' Association.

POOR TEST RESULTS LEAD TO MATH CONSORTIUM

10-state partnership is using results from an international test to address deficiencies in U.S. 8th graders' mathematics skills.

The Mathematics Achievement Partnership is responding to weaknesses in middle school math performance exposed by the Third International Mathematics and Science Study (TIMSS). Ten states (see box below) and the organization Achieve, Inc. are identifying instructional materials and professional development to help students and teachers prepare for a rigorous 8th-grade assessment that the partnership will design. The math teaching aids and training will be made available to states this spring for use next school year. The completed assessment will be available in spring 2002.

> "The bulk of what is asked on our 8th-grade math tests is content that other countries expect 5th and 6th graders or younger students to master."
>
> — William Schmidt, TIMMS national research coordinator



Participating States

Illinois
Indiana
Maryland
Massachusetts
Michigan
New Hampshire
North Carolina
Vermont
Washington
Wisconsin

Lessons from TIMSS

While most states test their 8th graders in math, there is concern about the rigor of their standards. In some states, 80% of the students meet standards; in others, a majority fails to do so. Despite progress separately on standards and testing, states have no way to compare results across their borders or against a common high benchmark. When compared to students in other countries, U.S. 8th graders perform below the international average.

Achieve asked William Schmidt, TIMMS' national research coordinator, to analyze the standards and assessments from the participating states based on the international study. He concluded that many states' 8th-grade tests concentrate largely on basic skills that other countries finish teaching in elementary school.

"While states make reference to them, the areas considered central to middle school math in the highest-achieving countries are not adequately measured by state tests. As a result, we can only speculate about whether that material is taught, and TIMSS gives us reason to believe it is not," Schmidt said. "The bulk of what is asked on our 8th-grade math tests is content that other countries expect 5th and 6th graders or younger students to master."

Foundations of higher math

The Mathematics Achievement Partnership will focus on the fundamental areas that form the core expectations in high-achieving countries: the underpinnings of algebra and geometry equations, formulas, two-dimensional geometry, measurement, proportionality, exponents, roots, radicals, slope, and congruence and similarity.

The initiative will provide states with tools to boost learning in these topics, as well as a common yardstick against which to measure progress. An internationally benchmarked assessment to be given near the end of 8th grade will inform parents, educators, employers and policymakers of how well students are mastering the foundations of algebra and geometry.

Achieve, Inc. is an independent, bipartisan, nonprofit organization formed to serve as a resource center to states on standards, assessment, accountability and technology. For more information, see the Achieve Web site, www.achieve.org/achieve/achievestart.nsf/ and click on The Mathematics Achievement Partnership.■

BEST COPY AVAILABLE

HOW STATES ARE RESPONDING TO LOW-PERFORMING SCHOOLS

by Katy Anthes, Susie Saavedra, Judie Mathers and Jane Armstrong

States increasingly are putting high-stakes accountability systems in place to assure all students meet high academic standards. Not all schools, however, are successful in helping students meet these standards and, as a result, may be designated by the state as low performing. With this designation, schools and districts may become eligible for assistance, or the state may apply sanctions.

States with accountability systems often intervene in such schools, districts or both. (The word "school" in this article applies to districts as well.)

How do states identify low-performing schools, and what happens to them?

Standards are statements of what students should know and/or be able to do, usually at each grade level. States measure school performance by student assessment results and other indicators, such as attendance and dropout rates. Schools whose students do not achieve at least "basic proficiency" on standards are placed in one of several categories, typically: watch/warning, probation or failing/in crisis.

As of December 1998, 35 states had statutes or regulations dictating specific sanctions for low-performing schools or districts. A school placed in one of these categories has a period of time to make specific improvements. If it doesn't, the state applies additional sanctions. The chart below summarizes the sanctions most commonly used.

Category	Sanction/Intervention
Watch/Warning	• Letter of notification • Require creation of school improvement plan (SIP) • Publicly reported list
Probation	• SIP implemented • Additional funds provided • Expert teacher assigned • Require use of comprehensive reform plan • State assistance team • Enrollment options provided
Failing/In Crisis (no improvement over time)	• Loss of accreditation • Reconstitution • Reorganization • Takeover — set up state-run charter or privately run school • School closure

Generally, a state accountability system includes a number of performance indicators publicly reported for each school. Some typical indicators include graduation rate, state assessment scores for all students or a sample of students, attendance rate and dropout rate.

Continued on next page

Continued from previous page

Watch/Warning

If a school falls short of meeting the state's basic performance goals, the state department of education notifies the district and school that the school has been placed on watch or warning status. This notification appears in the local and state media, school and district report cards, and parent letters.

In this category, the school is provided the opportunity to meet goals not previously met within a specified time. The school must complete a school improvement plan and is expected to become involved in the process. With the designation of "watch/warning," schools are expected to make improvements with little or no funding assistance. If sufficient improvements are not made within the specified time, the school may be designated as on probation.

"In recent years, an increasing number of state takeovers have resulted in authority being shifted to non-education leaders such as the governor or mayors."

Probation

At this level, the state may hire specialists for school improvement and assign an expert teacher to the school. Schools in this category also may be required to implement a comprehensive reform plan and/or give families the option of moving their children to other schools. For example, Illinois provides a probation manager and external partner to assist in developing and implementing an improvement plan. In Kentucky, the school is assigned a regional school support team and a "Highly Skilled Educator" and becomes eligible for school improvement funds.

As of January 1999, seven states (Texas, Oklahoma, Louisiana, Kentucky, North Carolina, West Virginia and New York) had legislation permitting students in low-performing schools to enroll in more successful ones.

Failing/In crisis schools

If the school still has not improved despite district and state assistance, it can be designated as "failing" or "in crisis." These schools require more drastic measures, such as reconstitution, takeover and closure.

During a school reconstitution, the state may replace principals, teachers and other staff and implement a new curriculum. Since the first reconstitution plan was implemented in San Francisco in 1983, at least six other states have reconstituted schools, including Colorado, Illinois, Maryland, New York, Ohio and Texas. States not having reconstitution statutes may sanction a school by removing accreditation. Ten states have legislative authority to remove the principal of a failing school. Sixteen states have authority to reconstitute, take over or close schools (see box below).

Reorganization within a district may take the form of appointing a new superintendent, reorganizing other administrative personnel, or appointing or requiring the election of a new school board. Traditional state takeovers of districts occur when the state legislature, board of education or the federal courts reassign district authority to the state department of education or another prescribed authoritative body.

In recent years, an increasing number of state takeovers have resulted in authority being shifted to non-education leaders such as the governor or mayors. For example, in 1995, the state legislature shifted control of the Chicago Public School system to the mayor who then was responsible for appointing a new school board and other district officers.

Sanctions and interventions may be helping schools improve. For example, in Florida's Miami-Dade School District, 45 schools implemented an intensive three-year corrective action plan, including schoolwide reading programs and improved technology. By the end of three years, all schools had made significant progress and were removed from the state's list of low-performing schools.

Anthes is a research assistant, Saavedra is an ECS intern and undergraduate student at the University of Denver, Mathers is a policy analyst, and Armstrong is director of policy studies for ECS.■

States having legislative authority to remove the principal of a failing school		States with legislative authority to reconstitute, take over or close schools		
Alabama	Michigan	Alabama	Michigan	Oklahoma
Delaware	Nevada	Illinois	Nevada	Rhode Island
Illinois	New York	Indiana	New Mexico	South Carolina
Kansas	North Carolina	Kansas	New York	Texas
Louisiana	South Carolina	Louisiana	North Carolina	Vermont
		Maryland		

PERFORMANCE MANAGEMENT, NOT JUST ACCOUNTABILITY

by Peter Robertson

Most of us have had this experience: After watching you do something badly (misassemble a child's bicycle, ruin a cake, etc.), someone will tell you what you should have done or, worse, merely what you did wrong. I'll call this person an "accountability expert," and most of us would like to throw him and his comments out a window.

What we prefer is some timely advice or observations ("Why don't you try putting the tire on the wheel before putting the wheel on the bike?" or, "That seems like a lot of flour....") from someone I'll call a "performance management" expert. We benefit more from input during the process that is designed to help us perform better. The same applies to the work of students, teachers and principals who, I believe, need performance management.

Accountability reflects history

Accountability — the ability to answer or account for — is about history, about what has been done. In the process-focused environment of public education, it often has been used in a "CYA" mode: "These are the steps we took, and here's what happened." Accountability often has obscured the question of what should have been done. And it has ignored the more important question of what must be done now.

Performance management, by contrast, is concerned with where we need to be and what it will take to get there. Attention is on the gap between current and needed outcomes and opportunities to do things better. Process, or what steps to take, is only dealt with in the context of what needs to be done. Historical data are useful only insofar as they help us identify the performance gap, motivate us to close it or inform us as to how to close it. Performance management data do not need to be perfect, only good enough to identify, motivate and inform.

Performance management in action

Here are some examples of how the Cleveland Municipal School District recently has used testing data for performance management.

Historically, we have published test data to tell the community what happened — an accountability model. This year, we invested considerable energy in formatting the test data, merging the data with new student assignment and attendance files, and giving every teacher a roster of their students' test scores and average daily attendance from last year. We are working to automate the process so that next year teachers and administrators can pull test scores and other data directly from a secure area of the district's Web site.

We also used test performance and analyses of student-specific test data to predict the performance of schools. We encouraged schools to use last year's data to target their intervention efforts with students. And, we created quarterly interim tests and used the data to help schools understand whether they are "on target" or need to reallocate teaching resources and student time to reach their goals.

Danger in the use

The danger of abuse is not in the numbers themselves; it is in how people use, or do not use, the numbers. "Accountability" makes insufficient use of the data available, and if performance management takes chances on less-than-perfect data in an attempt to make it useful, that is the greater good.

It would, of course, be evil and tragic to use performance-management data to track or discard children not predicted to be successful, or to use it to set unreasonable goals for teachers or administrators. But that kind of abuse is not the fault of the data; it is the fault of bad management and an unproductive work environment.

Performance management data, by focusing on what is possible in concrete, measurable terms, will spotlight those problems much faster than years of accountability data. Given how little awareness of those problems accountability has generated over the years, it's certainly worth a shot.

Robertson is executive director of the Office of Research, Evaluation, and Assessment of the Cleveland Municipal School District.■

"Performance management, by contrast, is concerned with where we need to be and what it will take to get there."

HIGH-STAKES TESTING PRESSURES TEACHER EDUCATION

by Michael Allen

High-stakes testing for students puts pressure on teachers to ensure their students do well on tests. This pressure increases dramatically when high stakes for adults — pay increases, job retention or school reconstitution — are attached to student test results. Under these circumstances, teachers are likely to focus significant attention on ensuring that their classroom activities and instruction prepare students for high-stakes examinations.

When high-stakes tests are closely tied to statewide student performance standards, preparing students for success on the tests should be accomplished by incorporating standards into the curriculum. Ideally, teachers would focus their teaching broadly on the general content knowledge and analytical skills the standards and tests are meant to reflect. This alignment between standards and curriculum is precisely what states seek when they develop content and performance standards and assess how well students meet the standards

Such alignment, however, is rare. While teachers are willing to integrate state standards into their curriculum, many need instruction in how to do so. The challenges facing teachers include: (1) knowing how to use the standards, (2) having adequate subject-matter knowledge that the standards require of students and (3) being tempted to ignore the standards and teach to the tests themselves (see page 7 for more on teaching to the test).

Teacher programs getting involved

Because the stakes attached to state tests are growing so high for both students and teachers, some teacher preparation programs — especially at state-supported universities — are moving to ensure that their curriculum reflects state student performance standards. These programs include several key components:

- ☐ Acquainting teacher candidates with the state standards system
- ☐ Requiring graduates to demonstrate content knowledge sufficient for them to address student content standards at the grade levels they will be teaching
- ☐ Ensuring that candidates learn, and know how to apply, content-based pedagogical and assessment practices associated with standards-based teaching
- ☐ Teaching candidates how to integrate student standards into their curriculum.

A few states have begun to require their teacher preparation programs to demonstrate that their graduates will be proficient in standards-based teaching. Many teacher preparation programs, however, have been slow to recognize the need to add the standards component into their program, and many of those that have done so are not as effective as they could be.

Professional development

For teachers already in the classroom, a number of states have recognized the importance of supporting professional development that helps teachers understand and integrate the student standards into their classroom. Some states and districts have developed model curricula to guide teachers in using the standards in their teaching; a few have online electronic support to increase teacher access to resources on the standards. Other states use their regional service centers to train practicing teachers to employ standards more effectively.

Nevertheless, state efforts in this area are often inadequate and underfunded. Moreover, professional development incentive structures — re-licensure or continuing certification — are generally silent about how teachers are to incorporate standards into curriculum.

The increase in high-stakes testing represents a pivotal area for the standards movement in states and is likely to survive or fail on teachers' ability to help students reach those standards. State leaders need to be sure postsecondary institutions are incorporating standards into teacher training programs and that current teachers have professional development opportunities that help them better integrate the standards into their classrooms.

Allen is an ECS policy analyst in charge of the quality teaching initiative. ☐

BEST COPY AVAILABLE

COORDINATED ASSESSMENT SYSTEMS A PLUS

by Edward D. Roeber

The current emphasis on assessment makes it prudent for states and local districts to develop a coordinated assessment system so that data collected at the state and local levels provide a fairly complete picture of student achievement.

A coordinated set of assessments is one in which assessments used at different levels of the education system fit together. This system eliminates redundant information, yet uses multiple sources to create a composite of student and school information.

State and local assessment systems can be built in tandem, based on a common set of content standards, to ensure the skills assessed are related and that different assessments work together. When assessment systems already have been developed at the state and/or local levels, coordination can occur in one of two ways. One level (e.g., the district) can use or adapt the assessment developed by the other level (e.g., the state), or the two levels can look for commonalities among their standards and assessments levels and report information derived from assessing those.

Assessment purposes

Coordinated assessment systems make sense because they also use available resources to collect information most useful for the decisions that need to be made at each education level. In addition, they reduce the number of "mixed messages" that local educators and the public receive about "what is important." By developing one set of content standards, with appropriate curricula and instructional strategies, the likelihood that students are taught the important skills also increases.

Assessment gaps

State and local education officials use large-scale assessment for various reasons. Student assessment is viewed as the means for setting higher, more rigorous standards for student learning, focusing staff development efforts for the nation's teachers, encouraging curriculum reform, and improving instruction and instructional materials in a variety of subject matters and disciplines. Assessment also

Continued on next page

> "A coordinated set of assessments is one in which assessments used at different levels of the education system fit together."

Continued from previous page

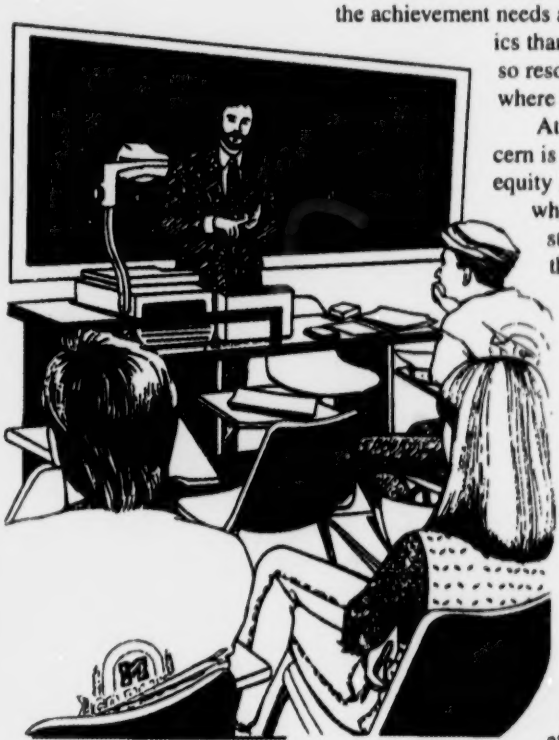
"This growing shift in emphasis in student outcomes is leading some at the national, state and local levels to emphasize new means of assessing student performance."

may serve to hold schools accountable for whether reforms have been effective.

Over the past decade, however, it has become clear that assessment programs that feature accountability for performance as a key purpose are often unable to fulfill the equally popular purpose of improving instruction. This is because accountability measures administered at the state level tend not to provide detailed information to teachers on a timely basis and because the information often does not assess students in a fashion most related to day-to-day instruction. The types of assessments most useful to teachers, though, do not often lend themselves to the public credibility demanded of accountability assessments.

Certainly, the parent's information needs are different from those of the teacher; the parent wants to know what his or her child can do and not do, while the teacher is more concerned with what additional work a student may need.

The building principal wants to know if achievement in the school is comparable to that elsewhere and, more broadly, whether students are learning what they need to learn. At the district level, the concern may be more whether the achievement needs are greater in mathematics than reading, for example, so resources can be allocated where most needed.



At the state level, the concern is often whether there is equity in school programs and whether students in the state are competitive with those in other states. This competitive concern also permeates the discussions at the national level where the underlying worry is about how much American students are learning in comparison to their peers in other countries.

Assessment design and format

These information needs, which may be very different at each level, often form the basis for assessment design. In top-down models, assessments that meet the needs of policymakers at the state or national levels are developed and implemented, with the presumption that the information will be useful to building principals, teachers and parents as well.

An emerging alternative to this is to build an assessment system that teachers, parents and students need and presume that users at the district, state and national levels can have their questions answered by aggregating the types of assessments used at the classroom levels.

A variety of content-area groups are reexamining what they view as important and how schools should be teaching these outcomes. A common element is the de-emphasis of content knowledge and an emerging emphasis on application and use of the content. This growing shift in emphasis in student outcomes is leading some at the national, state and local levels to emphasize new means of assessing student performance, such as portfolios, projects, exhibitions, demonstrations, individual performance assessments, group performance assessments and hands-on assessments.

Questions about strategies

Yet, in recent years, questions have been raised about the feasibility of using such innovative assessment strategies on a widescale basis. Issues of assessment time, generalizability, quality and breadth of resultant information, and costs have emerged as major impediments to the adoption of performance assessment in many large-scale assessment programs. Policymakers and others view these instructional-related assessment strategies as the ones to use for assessment programs tied to instructional improvement, however.

Each of the 45 states that have some form of large-scale assessment program has a different configuration of grades and subject areas assessed. They use different forms and, in some cases, multiple forms of assessment. Therefore, each state's assessment system could look different. See page 19 for an example of how such a system could be developed. (It is not intended to serve as a model coordinated assessment system.)

Roeber is vice president, external relations, Advanced Systems in Measurement & Evaluation in Dover, New Hampshire. This article is adapted from a paper he wrote while at the Council of Chief State School Officers.□

An Example of a Coordinated Assessment System

1. The state develops a set of content standards in selected areas with local district input. Most school districts adopt the state standards as their own.
2. In each area, the state coordination team develops an assessment blueprint describing the manner in which the content standards are to be assessed at the state, district and classroom levels.
3. The state selects subjects for statewide assessments to be administered in certain grades. The purpose of the assessments is primarily to hold schools accountable for student performance. Results are reported to parents, teachers, schools and districts.
4. Performance standards are created for each area in which the state has created content standards. These standards ensure assessments can be used to judge the performance of students and schools.
5. For each area in which the state has developed content standards, the state coordination team also develops a professional development program to ensure that all local educators are able to address the content standards and help students achieve at high levels.
6. The state creates the assessments that will be used, with the state coordination team overseeing the work to assure the assessments match the content standards and fulfill the purposes of the overall assessment system.
7. The state creates other assessments (portfolio assessments, performance events, performance tasks, plus more conventional selected-response and open-ended assessments) for use as "off-grades" throughout the school year. These assessments provide information teachers can use to improve the learning of individual students, as well as group information to improve the instructional program at the school and classroom levels.
8. The state sees that the assessments are created, validated and distributed across the state. As part of this process, the state administers the assessments to a sample of students statewide at each grade level, develops scoring rubrics and training materials for each open-ended or performance measure, and prepares the materials for distribution to school districts.
9. Assessments are tried out in a representative set of classrooms around the state with the results used in several ways: to refine the assessments themselves, to refine the assessment administration directions, and to revise and expand the scoring rubrics.
10. The state provides ongoing information and professional development opportunities to all local school districts. Assessment information collected by classroom teachers is summarized at the building level. District and school summaries are added to provide a more complete picture of student achievement.

WHAT MAKES A GOOD HIGH-STAKES ACCOUNTABILITY SYSTEM?

In spite of the controversy surrounding standards-based assessment systems, policymakers can take steps to alleviate problems and improve the impact and uses of assessment systems. The National Center for Research on Evaluation, Standards, and Student Testing has these suggestions:

1. *Set standards that are high, but attainable.* Standards that are too low or too high cause the public to lose faith in public schools or believe they are beyond improvement.
2. *Develop standards first, then assessments.* Imposing performance standards on existing tests doesn't work.
3. *Include all students in the testing program except those with the most severe disabilities. Use "accommodated" tests for students who do not speak English or whose disabilities require it.* Report scores by subgroup to provide accurate and useful information on student and school progress.
4. *Use new high-quality assessments each year that are comparable to those of the previous year.* Reusing the same test from year to year is likely to lead to distorted results, such as inflated test scores, or issues such as narrow teaching to the test.
5. *Don't rely solely on a single test when making important decisions about students.* Use multiple indicators such as grades, attendance, Advanced Placement course enrollment, performance assessments, etc. when making decisions about promotion, retention, graduation or rewards.
6. *Place more emphasis on comparisons of performance from year to year than from school to school.* This recognizes that schools start in different places but maintains an expectation of improvement for all.
7. *Set both long- and short-term goals for all schools to reach.* Short-term goals allow schools to start in different positions. Long-term goals permit high expectations for all schools, with a subsequent expectation that lower-achieving schools will have greater growth rates than high-achieving schools.
8. *Report uncertainty about the testing system.* Like an opinion poll, there is uncertainty in any education testing system that should be reported in all test results.
9. *Evaluate unintended negative effects of the testing system, as well as hoped-for effects.*
10. *Improve the education system as a whole; don't just add more testing or new testing systems.* Narrowing the achievement gap means children must have the teachers and resources they need to reach high expectations.□

Check out our Web site to see what's new at ECS.

www.ecs.org

EDUCATION COMMISSION OF THE STATES

ECS

Education Commission of the States
707 17th Street, Suite 2700
Denver, Colorado 80202-3427

Forwarding Service Requested

Nonprofit Org.
U.S. POSTAGE
PAID
Denver, Colorado
Permit No. 153